

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2001, 04:04:14 ; Search time 6128.19 Seconds

(without alignments)  
2020.032 Million cell updates/sec

Title: US-09-526-329-38

Perfect score: 1152  
1 gagcgcggtccggttgcag.....aaaaaaaaaaaaaaaa 1152

Scoring table: IDENTITY\_NDC  
Gapop 10.0, Gapept 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlom:\*  
5: em\_estlpi:\*  
6: em\_estlba:\*  
7: em\_estlro:\*  
8: em\_estrov:\*  
9: em\_hic:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pln:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rnd:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	664.4	57.7	674	10	AM978572 EST390681
2	606.6	52.7	657	10	AM370369 RCL-BT025
3	498.2	43.2	508	10	AA825207 oc66607.s
4	493	42.8	495	10	AM575965 UI-HF-BL0
5	454.4	39.4	456	10	AI379831 t629409.x
6	407	35.3	407	10	AI224367
7	396.4	34.4	398	10	AM292092
8	374.4	32.5	376	10	AI832007
9	372.8	32.4	525	10	AI250843
10	349.4	30.3	358	10	AM404913
11	331.4	28.8	338	10	AA825205
12	312.8	27.2	328	10	AM003249

Result No.	Score	Query Match	Length	ID	Description
13	304	26.4	315	10	AI434322
14	283.4	24.6	300	10	AI250350
15	262.8	22.8	492	13	AO589625
16	261.8	22.7	265	10	AM768936
17	239.2	20.8	463	13	AO572839
18	189	16.4	389	10	AM480639
19	173.4	15.1	560	10	BE685287
20	169.4	14.7	457	13	AO518451
21	166	14.4	178	10	AI468861
22	143.2	12.4	148	10	AA689513
23	125.8	10.9	161	11	BF082894
24	119	10.3	282	10	AM345961
25	93.8	8.1	472	11	BF605796
26	73	6.3	365	10	AM014278
27	63.6	5.5	537	10	AI391250
28	62.8	5.5	431	10	AM405686
29	62.8	5.5	449	11	BF128815
30	62.8	5.5	493	10	BE267581
31	62.8	5.5	494	10	AM500764
32	62.8	5.5	494	10	AM732980
33	62.8	5.5	511	10	BE398017
34	62.8	5.5	515	10	BE267867
35	62.8	5.5	535	10	BE268990
36	62.8	5.5	547	10	BE269500
37	62.8	5.5	553	10	BE513898
38	62.8	5.5	554	11	BI225045
39	62.8	5.5	554	11	BI225862
40	62.8	5.5	559	11	BI261356
41	62.8	5.5	566	10	BE295457
42	62.8	5.5	570	11	BI226934
43	62.8	5.5	574	10	AM732867
44	62.8	5.5	576	10	AM328037
45	62.8	5.5	578	11	BG759756

#### ALIGNMENTS

RESULT 1  
AM978572/c  
LOCUS AM978572 674 bp mRNA  
DEFINITION EST390681 MAGE resequences, MAGP Homo sapiens CDNA, mRNA sequence.  
ACCESSION AM978572  
VERSION AM978572.1 GI:8169840  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 674)  
Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt  
,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
Unpublished (2000)  
CONTACT: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnqu@tigr.org  
Plate: 394  
Seq primer: Forward.  
FEATURES  
SOURCE  
1. 674  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGP"  
/note="Vector: pBluescriptKm"  
BASE COUNT 185 a 178 c 191 g 120 t  
ORIGIN



Db 57 GACCTCACAGTACCTGTTAATTAACCTTATTGCTC 21  
 RESULT 3  
 AA825207/c 508 bp mRNA EST 21-APR-1998  
 LOCUS oc66e7.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:1354692 3',  
 DEFINITION mRNA sequence.  
 ACCESSION AA825207  
 VERSION AA825207.1 GI:2898504  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 508)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
 Ph.D., Gerald Marti, M.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdnp/image/image.html  
 Insert Length: 1089 Std Error: 0.00  
 Seq primer: 40m13 fwd. ET from Amersham  
 High quality sequence stop: 449.  
 Location/Qualifiers  
 1..508  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1354692"  
 /clone\_lib="NCI\_CGAP\_GCB1"  
 /tissue\_type="germlinal center B cell"  
 /lab\_host="DH10B"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD04+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 15'-TGTTACCATCTGAGTGGAGGCGCCCTCATTTTCTTTTCTTTT-3'  
 1 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 125 a 129 c 132 g 122 t  
 ORIGIN  
 Query Match 43.2%; Score 498.2; DB 10; Length 508;  
 Best Local Similarity 99.4%; Pred. No. 2; 4e-81;  
 Matches 500; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 650 actggggccctgtctgtctgagccagcttccctgtctgtctgagctgtgtct 709  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 508 ACTGGGGCCCTGTCTGTCTGAGCCAGTTTCCCTGCTGGTGCAGGTGGGTTCT 449  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 710 tttccctctgtgcccctatctatcttctatgagcaccatcccaatccctatacc 769  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 448 TTTCTTCTGTGSCCTCTATCTTCTTCTATGACACCTCCCAATCCCTTATACC 389  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 770 caccagatgtgtgcccagcagctcaccagaccctccagtcagctgtgtgtaaac 829

Db 388 CACCAGATGTGTGCCAGCCAGGCTCCAGCACCAGCCAGGCTGATTTGGAAC 329  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 830 ttaccatcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 889  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 328 TCACCATGCGCAGGCGAGTGTGCTGTTAAGATGATTCACGAGGACCCAGTTCGA 269  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 890 tctgacttgatgacctgtgtgtatcagcttctgtctgacacttggccgaatagacc 949  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 268 TGTGAGTGTGATGCCCTGTGTGATGATGATGATGATGATGATGATGATGAT 209  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 950 agtctgacgaacatgtacacgagcagcagcagcagcagcagcagcagcagcagc 1009  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 208 AGTCTGACGACACCATGTACACCGAGCCTCAGTGCATCTGCACAGTGGGAGCA 149  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1010 tggagggatgggt 1069  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 148 TGGAGGAGTGGGTTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 89  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1070 gctagtcctggaagtgagcctcagcagcagcagcagcagcagcagcagcagcag 1129  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 88 GCTAGTCCGTGAAGGAGGACCTCAGACTGCTTAACTTAACTTATGCTACTGTCA 29  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 1130 aaaaaaataaaaaaataaaaaa 1152  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 28 AAAAAAAAAAAAAAAAAAAAAA 6  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 RESULT 4  
 AM575965/c 495 bp mRNA EST 15-MAR-2000  
 LOCUS UT-HF-BL0-3cn-h-03-0-UT-s1 NIH\_MGC\_37 Homo sapiens cDNA clone  
 DEFINITION IMAGE:3059932 3', mRNA sequence.  
 ACCESSION AM575965  
 VERSION AM575965.1 GI:7247420  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 495)  
 NIH-MGC http://mgc.ncbi.nlm.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 The sequence contained an oligo-dT track that was present in the  
 oligonucleotide that was used to prime the synthesis of first  
 strand cDNA and therefore this may represent a bonafide poly A  
 tail. Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M.B. Soares Lab  
 cDNA Library Arrayed by: M.B. Soares Lab  
 DNA Sequencing by: M.B. Soares Lab  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bdnp/image/image.html  
 Seq primer: M13 forward  
 POLYA=Yes.  
 Location/Qualifiers  
 1..495  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3059932"  
 /clone\_lib="NIH\_MGC\_37"  
 /tissue\_type="lymph"  
 /cell\_type="germlinal center B cells"  
 /cell\_line="MGC85"  
 /lab\_host="DH10B (UT)"  
 /note="Vector: p773-Pac; Site\_1: NotI; Site\_2: Eco RI;  
 Constructed from size fractionated cytoplasmic mRNA  
 (1.5-2.5kb). Directionally cloned. Cells provided by Louis  
 M. Staudt, Ph.D. Library Preparation by Maria de Fatima  
 Bonaldo, Ph.D. and M. Bento Soares, Ph.D."



COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Unknown library type  
Insert Length: 675 Std Error: 0.00  
Seq primer: -40up from Gdbco  
High quality sequence stop: 383.  
Location/Qualifiers  
1..407  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="2000485"  
/clone\_id="NCI\_CGAP\_Lym12"  
/tissue\_type="lymphoma, follicular mixed small and large cell"  
/lab\_host="DH10B"  
/note="Organ: Lymph node; Vector: pCMV-SPORT6, Site: 1;  
Site: 2; Not: 1; Cloned unidirectionally. Primer:  
Oligo dt. Average insert size 1.25 kb. Life Technologies  
catalog #: 11547-015"

BASE COUNT 100 a 109 c 107 g 91 t  
ORIGIN

Query Match 35.3%; Score 407; DB 10; Length 407;  
Best Local Similarity 100.0%; Pred. No. 1.1e-64;  
Matches 407; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 725 ctaatgctatctctatagatgacacccaacacccctcaccacacagagatgtgtc 784  
|||||  
DB 407 CTCATGCTGATCTCTAGATGCTCAGTCCCAATCCCTTACATCCACAGAGATGTGTG 348  
|||||  
OY 785 ccaagcagagcctcagacacacacacagctcagctcagatgaaacacacacagcagc 844  
|||||  
DB 347 CCACCCAGGCTCCAGACACCCCGAGCTCGATGTAAGAACCTCACCATGCGAGCG 288  
|||||  
OY 845 agtgctggtttaagagatgagagagagagagagagagagagagagagagagag 904  
|||||  
DB 287 AGTGGTTCGGTTTAAAGATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 228  
|||||  
OY 905 ccttggtgatactcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 964  
|||||  
DB 227 CCTGTGGTATCAGTTCTGCTGAGACACTTGGCCCAATTAATACATCTCAGTACGAC 168  
|||||  
OY 965 atgacacag 1024  
|||||  
DB 167 ATGACACAGGAGCTCTGAGTACATCTGACAGTGGAGAGATGAGAGAGAGAG 108  
|||||  
OY 1025 ggcctgtct 1084  
|||||  
DB 107 GGCCTGTCT 48  
|||||  
OY 1085 tgaacctcaagatcgtgttaataaactttatctcactgtcaca 1131  
|||||  
DB 47 TGACCTCAAGATGAGTGAATTAACCTTATGCTCAGTCAAAA 1

RESULT 7  
AM292092 398 bp mRNA EST 16-JAN-2000  
LOCUS UI-H-B12-agx-a-12-0-UI.s1 NCI\_CGAP\_Sub4 Homo sapiens cDNA clone  
DEFINITION IMAGE:2725553, mRNA sequence.  
ACCESSION AM292092  
VERSION AM292092.1 GI:6698728  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 398)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
JOURNAL Tumor Gene Index  
Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
The sequence contained an oligo-dt track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/ILNI, at:  
[www.bio.lnl.gov/bbrp/image/image.html](http://www.bio.lnl.gov/bbrp/image/image.html)  
Seq primer: M13 Forward  
POLYA=yes

FEATURES  
source Location/Qualifiers  
1..398  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="2725553"  
/clone\_id="NCI\_CGAP\_Sub4"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pRT3D-Pac (Pharmacia) with a modified  
polylinker. Site: 1: Not I; Site: 2: Eco RI; The  
NCI-CGAP\_Sub4 library is a subtracted library derived from  
the NCI-CGAP\_Sub2 library which is a subtracted library  
derived from the NCI-CGAP\_Sub1 library, which is a  
subtracted library derived from BI. BI constitutes a  
mixture of 21 normalized or subtracted NCI-CGAP  
libraries: NCI-CGAP\_C04, NCI-CGAP\_Pr22, NCI-CGAP\_Pr28,  
NCI-CGAP\_C010, NCI-CGAP\_C016, NCI-CGAP\_Kid5,  
NCI-CGAP\_Kid12, NCI-CGAP\_Kid3, NCI-CGAP\_Kid11,  
NCI-CGAP\_Lym2, NCI-CGAP\_Pr2, NCI-CGAP\_Kid11,  
NCI-CGAP\_Ly12, NCI-CGAP\_Brn23, NCI-CGAP\_C111,  
NCI-CGAP\_Lu24, NCI-CGAP\_Lu19, NCI-CGAP\_G04, NCI-CGAP\_G06,  
NCI-CGAP\_Brn25. These 21 libraries were pooled and a  
single-stranded DNA preparation of the resulting mixture  
was used as a tracer in a subtractive hybridization with  
a driver whose composition is detailed below:  
NCI-CGAP\_Kid3 pool 1: LLM 3334-3337, 3682-3683,  
3798-3803 (IMAGE cloneids 1322376-1323911,  
1456008-1456775, 1500552-1502855) NCI-CGAP\_Kid5 pool 1:  
LLM 3338-3342, 3722-3725, 3776-3778 (IMAGE cloneids  
132912-132831, 1471368-1472903, 1492104-1493255)  
NCI-CGAP\_Lu5 pool 1: LLM 3575-3582, 3851-3854 (IMAGE  
cloneids 1414920-1417991, 1520904-1522439) NCI-CGAP\_G04  
pool 1: LLM 3164-3167, 3716-3720, 3733-3735 (IMAGE  
cloneids 1257096-1258631, 1469064-1470983, 1475592-1476743)  
NCI-CGAP\_Pr22 pool 1: LLM 2457-2459, 2758-2759,  
3062-3068 (IMAGE cloneids 985608-986759, 1101192-1101959,  
121928-1220615) NCI-CGAP\_C010 pool 1: LLM 2644-2653,  
2871-2872 (IMAGE cloneids 1057416-1061255, 114584-1145351)  
Subtraction was performed as previously described  
(Bonaldo, Lennon & Soares (1996): Normalization and  
Subtraction: Two Approaches To Facilitate Gene Discovery.  
Genome Research 6, 791-806.)  
TAG\_LIB=NCI\_CGAP\_Lu5  
TAG\_TISSUE=Lung  
TAG\_SEQ=CAAC"

BASE COUNT 92 a 104 c 102 g 100 t  
ORIGIN

Query Match 34.4%; Score 396.4; DB 10; Length 398;  
Best Local Similarity 99.7%; Pred. No. 9.3e-63;  
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 747 cactcccaatccctcaccacacacacacacacacacacacacacacacacacac 806  
|||||  
DB 398 CACTCCCAATCCCTTCAATCCACACAGATGTGTGCTCCAGCAGGCTCTCAACACCC 339  
|||||  
OY 807 cagtcagctcgtgatgagaaacacacacacacacacacacacacacacacacacac 866  
|||||  
DB 338 CAGTCAGCTCGTGTATGGAACCTCACATCGCAGCAGATGCTTGAAGAGATG 279  
|||||  
OY 867 cattagagagagac 926  
|||||

Db 278 CATTACAGGAGCCAGCTGATGTGACTTGATGCCCTGTGGGTATCAATTCCTG 219

Qy 927 acacttggcccgaaatagatcagctgctgagcaagaatgtaacacggagcccaatga 986

Db 218 ACACCTTGGCCCGAAATAGATCCAGTGTGACAGCAAGCAATGTACACGGAGCTCAGTGA 159

Qy 987 gccacatcgacagctggagagcatggaaggatgggttggccctgctctgcttattca 1046

Db 158 GCCCATCTCCACAGTGGGAGCATGAGAGGATGGGTTGGCCCTGTCTCTCTTATTCA 99

Qy 1047 gtccttcagctcaaggaagatgctagccgctgaagtgtaacctcacagtaactgtaact 1106

Db 98 GTCCTTACCTCAGCAGGAGGATGTAGTCCGTGAAGGTACCTCAGTACTGTTAAT 39

Qy 1107 taactttatgtctactgtctcaaaaaaataaaaaa 1144

Db 38 TAAACTTTATGTCTCAGTCAAAAAAATAAAAAA 1

RESULT 8  
AI832007 376 bp mRNA EST 21-DEC-1999  
LOCUS AI832007/c wj99a02.x1 NCI\_CGAP\_Lym12 Homo sapiens cDNA clone IMAGE:2410922 3',  
DEFINITION mRNA sequence.  
ACCESSION AI832007  
VERSION AI832007.1 GI:5452678  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 376)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Life Technologies catalog #: 11547-015  
DNA sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLN at:  
www.bio.lnli.gov/DBP/IMAGE/IMAGE.html  
Insert Length: 1100 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 375.

FEATURES  
source  
1..376  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2410922"  
/clone\_1ib="NCI\_CGAP\_Lym12"  
/tissue\_type="lymphoma, follicular mixed small and large cell"  
/lab\_host="DH10B"  
/note="Organ: lymph node; Vector: PCMV-SPORT6; Site: 1;  
Sali; Site: 2; NotI; Cloned unidirectionally. Primer:  
Oligo dT. Average insert size 1.25 kb. Life Technologies  
catalog #: 11547-015"  
BASE COUNT 92 a 105 c 97 g 82 t  
ORIGIN

Query Match 32.5%; Score 374.4; DB 10; Length 376;  
Best Local Similarity 99.7%; Pred. No. 9.7e-59;  
Matches 375; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 753 caatcccttataccacacagatgtgtgcccagcagccttcagacaccccaagtc 812

Db 376 CAAATCCCTTCAATACCCACAGATGTGTGCCACGACGCTCCACGACCCCAAGTGC 317

Qy 813 agctcgatgtgaactcacacagcagcagcagctggttcggtttaagagatgcatatg 872

Db 316 AGCTCGATTTGGAACCTACCATCGCAGGCGAGGCTGTTAAGATGTGCATTAG 257

Qy 873 agggagcccaagctgagatgctgagacttgatgacctgtggtatcagttctgtaacat 932

Db 256 AGGAGCCACAGTCTGATGTGTGACTGTGATGCCCTGTGGGTATCAGTTCTCTACACAT 197

Qy 933 tggccggaatagatcagctgctgagcagaatgtaacccggaacccagtaagccca 992

Db 196 TGCCCGCAATATGATTCAGTGTGACGACCAAGTATACACGAGCCTCAGTACCAT 137

Qy 993 ctgacagtgaggagacatggaaggatgggttggccctgctctgcttattcaact 1052

Db 136 CTCACAGTGGGAGCATGAGGAGATGGGTTGGCCCTGTCTCTTATTCACTCT 77

Qy 1053 cagctcaggaaggaatgctagctgctgaagtgtaacctcacagtaactgtaact 1112

Db 76 CAGCTCAGGAGGATGCTAGTCCGTGAAGGTACCTCAGTACTGTTAATTAACT 17

Qy 1113 ttatgtctactgtca 1128

Db 16 TTATGTCTCAGTCA 1

RESULT 9  
AI250843 525 bp mRNA EST 21-DEC-1998  
LOCUS AI250843/c qx10a04.x1 NCI\_CGAP\_Lym12 Homo sapiens cDNA clone IMAGE:2000910 3',  
DEFINITION mRNA sequence.  
ACCESSION AI250843  
VERSION AI250843.1 GI:3847372  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 525)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
unknown library type  
Insert Length: 1149 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 336.

FEATURES  
source  
1..525  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2000910"  
/clone\_1ib="NCI\_CGAP\_Lym12"  
/tissue\_type="lymphoma, follicular mixed small and large cell"  
/lab\_host="DH10B"  
/note="Organ: lymph node; Vector: PCMV-SPORT6; Site: 1;  
Sali; Site: 2; NotI; Cloned unidirectionally. Primer:  
Oligo dT. Average insert size 1.25 kb. Life Technologies  
catalog #: 11547-015"  
BASE COUNT 143 a 141 c 136 g 104 t 1 others  
ORIGIN

Query Match 32.4%; Score 372.8; DB 10; Length 525;  
Best Local Similarity 99.3%; Pred. No. 1.7e-58;  
Matches 374; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 752 ccaatcccttataccacacagatgtgtgcccagcagccttcagacaccccaagtc 811

Db 377 CAAATCCCTTCAATACCCACAGATGTGTGCCACGACGCTCCACGACCCCAAGTGC 318

Qy 812 cagctcgatgtgaactcacacagcagcagcagctggttcggtttaagagatgcatatg 871







Query Match	26.48;	Score 304;	DB 10;	Length 315;
Best Local Similarity	99.7%;	Pred. No. 6.9e-46;		
Matches 315;	Conservative	0;	Mismatches 0;	Indels 1;
				Gaps 1

OY	812	caactcgtgatgttgaacaccacacatggacgaggaugtggttcgttttaagaagatggcatla	871
Db	315	CACTCTGATGATGGAACCT - ACCATCGCAGGCGAGTGGCTCGGTTTAAAGATGGCAATT	257
OY	872	gaggaagcccaagtcctgatalgtgaacttgatgacctgfgatlaacagttcagtcgaacat	931
Db	256	GAGGAGAGCCCACTGCTTGGATGTGACTTGGATGGCCCTGTTGGGTATCATGTTCTGTGACACT	197
OY	932	ttggaccgaatctgatccaaagtgcagtgagccaagaagatgtccaccggagacctcaatgaacca	991
Db	196	TGGCGCCGAATTGATTCACAGTCTCTGACCAACGAAATGTACACCGAAGCTTCAGTGAAGCCA	137
OY	992	ttctgacaagtggggagacatgtgaaggatlgagtttggcctgtgctcttcgcattatcaagcct	1051
Db	136	TCTGACACAGTGGGAGACACAGAGAGGAGGGGTTTGGCCCTGTGCTTGCATTATCAGTCCT	77
OY	1052	tcaagtcacggaaggaatgctcagtcacgttaagaagttaactaaagtaactgaactgtaataaac	1111
Db	76	TCAAGTCACGGAAGGAGTCAATGCTGCTGAAGGTACTACACGTAAGTGAATTAATTAAAC	17
OY	1112	tttatctgtacatgctc 1127	
Db	16	TTTTATGCTCATGTC 1	

RESULT	14
A1250350/c	
LOCUS	
DEFINITION	A1250350 300 bp mRNA EST 21-DEC-1998
ACCESSION	gq4ccl.xl NCI-CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2000352 3'
VERSION	A1250350
KEYWORDS	A1250350.1 GI:3846879
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 300) NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov unknown library type Insert Length: 441 Std Error: 0.00 Seq Primer: ~400P from Glibco High quality sequence stop: 269. Location/Qualifiers 1..300
JOURNAL	
COMMENT	
FEATURES	
source	

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:2000352"  
/clone_lib="NCI_CGAP_Lym12"
```

Query Match	24.6%;	Score 283.4;	DB 10;	Length 300;
Best Local Similarity	99.6%;	Pred. No. 4e-42;		
Matches 284; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

QY	845	aagtgctcgtttaaagatgacgtgaggaagccgcgtctgattctgagcttgatgc	904
Db	300	AAGTGCTCGTTTAAAGATGACGTGAGGAGGCCGCGTCTGATTCTGAGCTTGATGC	241
QY	905	ccgtgaggatcagctctgcgcgaacttggccggaattgatccagtgcttgagcaaga	964
Db	240	CCGTGAGGATCAGCTCTGCGCGAACCTTTGGCCCAATATGATCCAGTCTGACCAAGA	181
QY	965	atgtacacgggaacccctcagtaagccacttcgaacagtggggagcatggagatggatt	1024
Db	180	ATGTACACGGGAACCCCTCAGTAAGCCATCTGCACAGTGGGACCAATGAGGAGGATGGATT	121
QY	1025	ggctggtgctctgacttattcaagtcccttaagctcaaggaaagagatgctgcgttaag	1084
Db	120	GGCTGCTGCTCTGACTTATTCAAGTCCCTTAAGCTCAAGGAGGAGATGCTGCCTGAAGG	61
QY	1085	taactcaacgactcgtgttaattaaacttattgctactgttcaa	1129
Db	60	TAACTCAACGACTCGTGTAAATTAAACTTATTGCTACTGTTCAA	16

RESULT	15
LOCUS	AQ589625
DEFINITION	AQ589625 492 bp DNA
ACCESSION	HS-2132.AL.D12.T7C CTR Approved Human Genomic Sperm Library D
VERSION	AQ589625
KEYWORDS	sapiens genomic clone Plate=2132 COL=23 Row=G, DNA sequence.
SOURCE	GSS.
	AQ589625.1 GI:5020678
	human.

ORGANISM	Homo sapiens
TAXONOMY	Animalia; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 492)
REFERENCE	Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE	99380589
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L

401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com)  
BAC end Web Server: <http://www.husc.washington.edu>  
plate: 2132 row: G column: 23  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 492.  
location/Qualifiers  
1..492  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

Fri Nov 30 11:03:08 2001

us-09-526-329-38.rst

Page 10

```

/clone="Plate-2132 Col=23 Row=G"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelbac11; BAC clones in
E-Coli DH10B"

```

BASE COUNT	103 a	148 c	124 g	112 t	5 others
ORIGIN					

Query Match	22.8%;	Score 262.8;	DB 13;	Length 492;
	75.18;	Prod No. 1 00-38;		

Best Local Similarity 1/5.15; Pieu. NO. 1.35 20;  
Matches 396; Conservative 0; Mismatches 87; Indels 44; Gaps 4;

Oy	61	ccctccctggccgctctgtgagatccagagccctgagactcagaaatgaagaaagggagagacc	120
Db	4	CNCNCACACTGTCTGTGGATCTGGAGACCTGGGATCTGTGAGATGACAAAGGGGGAGACC	63
Oy	121	tggatgtactgtgtgtcgtgcggttcaatccctcgtgttaggaatggggcaag-----ggc	174
Db	64	TGCTGTGACTGTGTGGTGTGGGCTCAAGTCCCTCTCACAGAGAAAGGCCAGGTGCTCAAGC	123
Oy	175	tccacgggcagacagataataaacccagatcacagtgtcaatttggcagatgagatgcat	234
Db	124	TCCACATGACACACATATATACCTAGATACAGTGCACAGTGTGATGGATGTGCAC	183
Oy	235	accggggagctaacctctccctccgacagatgacctctccacagctccgcgcgtgtggag	294
Db	184	ACCCGGGAAACCACTGTCCCCAGCCAGTGGCTTCTCCACCTGTCCCGCTGGGCAG	243
Oy	295	ctctaacccgggagaaagtaaccgagccagggagattcagattcttgagaaatagcaaacat	354
Db	244	CTCTACCCCAAGAGAGAGATCTCAACACACAAATTCACAGTTCTGGAAATAGTGA-----	299
Oy	355	ggccagatgtactctatlgagcaagctgtgtcctaataatcagcggggagggaaagactga	414
Db	300	--CCAATATGCTCTCAATGAGAGCAAGCTGTCTTAACATATGGCCGCGAAGAGATTAACGA	357
Oy	415	caatggagatgtgcctgacctgtgacccctgtctctcttcttggcctgtgtctctcatgccc	474
Db	358	-----CTCTCTGTGGCTCATATCTCTTCCTTCNGGGCC	386
Oy	475	ctctcagtgagagatctctatgactaccgtctctctgtttgaaaccacagacagactctctt	534
Db	387	NCATCATGTAGATCTTTGTGTACTATTGCTATTCCTCCGCTTGACACACCACAGCGGCTCTTGG	446
Oy	535	caggcagaagcagatagggccctctcaacatccagtttctctctggttt	581
Db	447	CAGGCAAGATGAGTGGGGGCCCTTGC-CATCTCAACTTCTCTATTTTT	492

Search completed: November 29, 2001, 05:49:45  
Job time: 6331 sec